

What is claimed is:

1. A method comprising:
generating a number of Information-Gain (IG)-Trees based on a memory-learning technique; and
extracting entity names and relations between entity names based on the IG-Trees.
2. The method of claim 1, wherein the number of IG-Trees is generated based on raw data that has been annotated.
3. The method of claim 2, wherein the number of IG-Trees is generated based on a number of features of the annotated data.
4. The method of claim 1, wherein the number of IG-Trees is selected from a group consisting of a person-name IG-Tree, an entity-name IG-Tree, a noun phrase IG-Tree and a relation IG-Tree.
5. A method comprising:
receiving annotated data;
parsing the annotated data;
extracting a number of training sets from the parsed annotated data; and
generating a number of Information Gain (IG)-Trees from the number of training sets.
6. The method of claim 5, further comprising segmenting the annotated data.

7. The method of claim 5, wherein the number of IG-Trees is selected from a group consisting of a person-name IG-Tree, an entity-name IG-Tree, a noun phrase IG-Tree and a relation IG-Tree.

8. The method of claim 7, further comprising extracting entity names from an input document based on the number of IG-Trees.

9. The method of claim 5, wherein the person-name (IG)-Tree is generating using memory-based learning.

10. The method of claim 5, wherein the number of IG-Trees is generated based on a number of features of the annotated data.

11. A method comprising:
generating a person-name Information Gain (IG)-Tree and a relation IG-Tree from annotated data;
tagging and partial parsing of an input document;
extracting names of persons within the input document using the person-name IG-tree;
extracting names of organizations within the input document;
extracting entity names that are not names of persons and organizations within the input document; and
extracting relations between the identified entity names using the relation-IG-tree.

12. The method of claim 11, further comprising:

extracting noun phrases within the input document using a noun-phrase (IG)-Tree generated from the annotated raw data; and

classifying the noun phrases extracted using an entity name IG-tree.

13. The method of claim 11, further comprising partial parsing of the input document based on the entity names and the noun phrases.

14. The method of claim 11, wherein the person-name (IG)-Tree is generating using memory-based learning.

15. The method of claim 11, wherein the number of IG-Trees is generated based on a number of features of the annotated data.

16. A machine-readable medium that provides instructions, which when executed by a machine, cause said machine to perform operations comprising:

generating a number of Information-Gain (IG)-Trees based on a memory-learning technique; and

extracting entity names and relations between entity names based on the IG-Trees.

17. The machine-readable medium of claim 16, wherein the number of IG-Trees is generated based on raw data that has been annotated.

18. The machine-readable medium of claim 17, wherein the number of IG-Trees is generated based on a number of features of the annotated data.

19. The machine-readable medium of claim 16, wherein the number of IG-Trees is selected from a group consisting of a person-name IG-Tree, an entity-name IG-Tree, a noun phrase IG-Tree and a relation IG-Tree.

20. A machine-readable medium that provides instructions, which when executed by a machine, cause said machine to perform operations comprising:

receiving annotated data;

parsing the annotated data;

extracting a number of training sets from the parsed annotated data; and

generating a number of Information Gain (IG)-Trees from the number of training sets.

21. The machine-readable medium of claim 20, further comprising segmenting the annotated data.

22. The machine-readable medium of claim 20, wherein the number of IG-Trees is selected from a group consisting of a person-name IG-Tree, an entity-name IG-Tree, a noun phrase IG-Tree and a relation IG-Tree.

23. The machine-readable medium of claim 22, further comprising extracting entity names from an input document based on the number of IG-Trees.

24. The machine-readable medium of claim 20, wherein the person-name (IG)-Tree is generating using memory-based learning.

25. The machine-readable medium of claim 20, wherein the number of IG-Trees is generated based on a number of features of the annotated data.

26. A machine-readable medium that provides instructions, which when executed by a machine, cause said machine to perform operations comprising:

generating a person-name Information Gain (IG)-Tree and a relation IG-Tree from annotated data;

tagging and partial parsing of an input document;

extracting names of persons within the input document using the person-name IG-tree;

extracting names of organizations within the input document;

extracting entity names that are not names of persons and organizations within the input document; and

extracting relations between the identified entity names using the relation-IG-tree.

27. The machine-readable medium of claim 26, further comprising:

extracting noun phrases within the input document using a noun-phrase (IG)-Tree generated from the annotated raw data; and

classifying the noun phrases extracted using an entity name IG-tree.

28. The machine-readable medium of claim 26, further comprising partial parsing of the input document based on the entity names and the noun phrases.

29. The machine-readable medium of claim 26, wherein the person-name (IG)-Tree is generating using memory-based learning.

30. The machine-readable medium of claim 26, wherein the number of IG-Trees is generated based on a number of features of the annotated data.